Hedging Crop Yields Against Weather Uncertainties—A Weather Derivative Perspective

Samuel Asante Gyamerah [1,*](ID), Philip Ngare [2] and Dennis Ikpe [3,4]

[1]   Institute for Basic Sciences, Technology, and Innovation, Pan African University, Nairobi, Kenya
[2]   School of Mathematics, University of Nairobi, Nairobi, Kenya
[3]   African Institute for Mathematical Sciences, Cape Town, South Africa
[4]   Department of Statistics and Probability, College of Natural Science, Michigan State University, 220 Trowbridge Rd, East Lansing, MI 48824, USA
*   Correspondence: saasgyam@gmail.com

check for updates

Abstract: The effects of weather on agriculture in recent years have become a major global concern. Hence, an effective weather risk management tool (i.e., weather derivatives) that can hedge crop yields against weather uncertainties is needed. However, most smallholder farmers and agricultural stakeholders are unwilling to pay for the price of weather derivatives (WD) because of the presence of basis risks (product-design and geographical) in the pricing models. To eliminate product-design basis risks, a machine learning ensemble technique was used to determine the relationship between maize yield and weather variables. The results revealed that the most significant weather variable that affected the yield of maize was average temperature. A mean-reverting model with a time-varying speed of mean reversion, seasonal mean, and local volatility that depended on the local average temperature was then proposed. The model was extended to a multi-dimensional model for different but correlated locations. Based on these average temperature models, pricing models for futures, options on futures, and basket futures for cumulative average temperature and growing degree-days are presented. Pricing futures on baskets reduces geographical basis risk, as buyers have the opportunity to select the most appropriate weather stations with their desired weight preference. With these pricing models, farmers and agricultural stakeholders can hedge their crops against the perils of extreme weather.

1. Introduction

Agriculture continues to be an important sector that contributes to Ghana's exports earnings, inputs for most manufacturing sectors, and revenue generation for majority of the population. The agriculture and agribusiness sector account for a significant share of the major economic activities in Ghana and a major source of income for most smallholder farmers. It is reported to account for about 70% of the labour force and more than 25% of the gross domestic product in Africa [1]. This makes agriculture one of the most important and largest sector in the development of the economies in Africa, of which Ghana is a member state. However, in Ghana, the sector continues to be controlled by primary production as a result of high weather variability and hydrological flows, especially in the Northern savannas [2]. The Northern savannas of Ghana have experienced perennial extreme flooding and droughts, both linked to extreme heat and temperatures [2]. These have contributed to crop failures and have consequently led to extensive impacts on the economic activities of most rural farmers.

Weather variables are difficult to mitigate, especially for smallholder farmers in most developing and under-developed countries, and have great effects on the farming activities of these farmers. For this reason, an effective and reliable risk management tool, weather derivative (WD), is needed to hedge farmers and stakeholders from the peril of weather uncertainties. By linking the payoffs to a fairly measured weather index (e.g., temperature, rainfall, humidity, and sunshine), a WD reduces or eliminates the disadvantages of traditional insurance, such as moral hazards and information asymmetry. Turvey [3] examined the pricing of weather derivatives in Ontario and contended that weather derivatives and weather insurance can be used as a form of agricultural risk management tools. Zong and Ender [4] developed a novel type of weather derivatives contract called a climatic zone-based growth degree-day contract. Their aim was to mitigate weather risk in the agricultural sector of mainland China by introducing new types of temperature indices. Even though several weather risk management tools have been recently introduced into the WD market for smallholder farmers in most countries around the world, their purchases have been lower than expected. Among the reasons causing the low purchase of WD and the unwillingness of farmers to pay for WD in most communities are the lack of capacity to design and determine the value of this insurance product and, most importantly, high basis risk (product-design and geographical) in the contract design and implementation [5]. Basis risk in WD market is defined as the difference between the actual loss and the actual payout of WD [6]. Using a different index rather than the actual index that affects a specific crop yield at a location can lead to a larger gap between the real exposure and the payoff (product-design basis risk). Product design basis risk can be mitigated if the appropriate weather observation is used to construct the index for WD. Hence, a complete evaluation of the relationship between historical weather and crop yield data is significant for an effective and reliable design of WD. Forecasting of crop yields and feature importance of different weather indices for crop yields are, therefore, principal components for an effective rate-making process in the insurance/derivatives market. Another form of basis risk–geographical basis risk occurs when there is a deviation of weather conditions at the measurement station of the weather derivative and the weather conditions at the location of the buyer [7]. Spicka and Hnilica [8] evaluated the effectiveness of weather derivatives as a revenue risk management tool by considering crop growing conditions in the Czech Republic. The authors concluded that high basis risk can significantly misrepresent the payoff of the contract. Musshoff et al. [9], in their research, concluded that hedging effectiveness for the agricultural sector using weather derivatives is controlled by the contract design. They categorized basis risk into local and geographical basis risk, and asserted that basis risk has a greater influence on the hedging effectiveness of weather derivatives.

In this study, geographical basis risk is mitigated by pricing futures on a temperature basket rather than a single index contract. This requires the determination of the correlation between the locations under study.

In the literature of weather derivative pricing, different methods have been proposed for pricing temperature-based weather derivatives. Among these methods are the indifference pricing approach, actuarial pricing methods, the equilibrium model approach, and incomplete market pricing models. The indifference pricing approach is a valuation method for weather derivatives which is founded on the arguments of expected utility. It is also referred to as the incomplete market pricing model, reservation price, or private valuation. This approach for pricing has been used in most traditional financial derivatives where the market is incomplete (see [10–12]). In the weather derivative market, Davis [13] used the marginal utility approach or "shadow price" method of mathematical economics to price weather derivatives. The basis for their pricing approach was centered on the concept that investors in the weather derivative market are not representative but experience distinct risks that are linked to the effect of weather on their business. Brockett et al. [14] used the indifference pricing approach to value weather derivative futures and options. The relationships between the indifference pricing and actuarial pricing approaches for weather derivatives were studied in a mean-variance context, where they provided instances that the actuarial pricing method does not give a distinct

valuation for weather derivatives. From the concept of utility maximization, Barrieu and Karoui [15] calculated the optimal profile (and its value) of derivatives written on an illiquid asset; for example, the weather or a catastrophic event. Cao and Wei [16] generalized the equilibrium model of [17] by incorporating weather as a basic variable in the economy. Based on this, they proposed an equilibrium valuation framework for temperature-based weather derivatives. Using a model that captured the daily temperature dynamics of five cities in the United States of America (Atlanta, Chicago, Dallas, New York, and Philadelphia), they performed numerical analysis for forward and option contracts on heating degree-days (HDDs) and cooling degree-days (CDDs). Their analysis revealed that the market price of risk is mostly trivial when linked to the temperature variable, particularly when the aggregate dividend process is mean-reverting. They showed that unrealistic assumptions in historical simulation methods result in inaccurate pricing of temperature-based weather derivatives. As stated by [18], the weather derivative market is a typical example of an incomplete market. Hence, pricing models based on incomplete markets are the most applicable valuation method for weather derivatives. These models consider the hedgeable and unhedgeable components of risk. The price of a weather derivative is usually dependent on different weather indices, such as HDD, CDD, Pacific Rim (PRIM), cumulative average temperature (CAT), and growing degree-days (GDD). Different authors [18,19] have used the HDD and CDD indices as the major indices for pricing weather derivatives for the energy industry.

The contributions made in this study are: (1) We are able to empirically determine the main underlying weather variable (average temperature) that affects the yield of the selected crop using machine learning ensemble techniques and feature selection through crop yield forecasting, rather than the usual assumption of using temperature as the underlying without proper empirical studies. This will eliminate product-design basis risk during pricing of the weather derivatives. (2) Previous studies have either used a piecewise constant volatility function or a seasonal volatility (e.g., [18,20,21]). However, our proposed model includes a local volatility which is able to capture the local variations of the daily average temperature. (3) Futures, options on futures, and basket futures (futures for multi-dimensional locations) on CAT and GDD are priced using the constructed daily average temperature model. These pricing models for CAT and GDD are the first of their kind in the literature. (4) The basket futures pricing will help in mitigating geographical basis risks in the weather derivative market.

2. Crop Yield–Weather Model and Feature Importance for Weather Derivatives

Different factors, such as condition of the soil, the variety of seedling, and type and amount of fertilizer applied to the soil affect the yield and productivity of crops. These factors can be controlled by the farmer and industry players. Weather variables, especially surface temperature, rainfall, and humidity, are the principal drivers of the differences in crop yields [22]. These weather variables directly affect the moisture content of the soil and the level of nutrients in the soil. The effect of the uncertainties in the pattern of weather, both between and within planting seasons, can affect the crop production and yield significantly. This has significantly affected the yield of most crops, causing economic and food security risks in most developing and under-developed countries. Maize is the most widely cultivated staple crop in the Northern savanna, and is the major source of income for about 45% of households there [23]. For this reason, the yield of maize was used as a proxy to determine the effect of the selected weather variables on crop yield. This weather variable can, then, be used as the underlying for weather derivatives pricing.

2.1. Machine Learning Ensemble Technique for Weather-Crop Yield Model

Machine learning (ML) is used in regression and classification to solve most of the problems that arise as a result of nonlinearity in the features and the response variable. In this study, ML ensemble classification algorithms are used to predict the possibility of an improved crop yield harvest or crop yield loss (a two-class (binary) problem). The choice of an optimal ML algorithm for prediction is a major factor to consider in any forecasting problem. In our case, the chosen ML technique should be

able to predict whether there will be an increase or decrease in crop yields for a period of years with a small margin of error. The target variable for the ensemble classifier is whether there will be a loss or increase in crop yields for the year-ahead harvest; that is,

- If the long-term average crop yield ($\bar{Y}$) is smaller than or equal to the present years crop yield (Y_t), then there is an increase in crop yield; else:
- If the long-term average crop yield ($\bar{Y}$) is greater than the present years crop yield (Y_t), then there is a decrease or loss in crop yield.

We assign labels to this classification: "0" for a decrease/loss in crop yields and "1" for an increase in crop yields,

$$y = \begin{cases} 1 & \text{if} \quad \bar{Y} \leq Y_t, \\ 0 & \text{if} \quad \bar{Y} > Y_t. \end{cases} \tag{1}$$

Humidity, sunlight, rainfall, minimum temperature (minT), maximum temperature (maxT), and average temperature (aveT) are the features (**X**) used in predicting the target variable (**Y**). Now, suppose a set of features $\mathbf{X} = \{x_i \in \mathbb{R}^n\}$ with associated labels $\mathbf{Y} = \{y_i \in Y, y_i \in (0,1)\}$, are given for a training data set $T = \{(x_i, y_i)\}$. In this way, we solve the supervised classification problem where the learning model N depends on D.

The yearly datasets (2000–2016) for maize yield Bole, Tamale, and Yendi were taken from the Statistics, Research, and Information Directorate (SRID) of the Ministry of Food and Agriculture, Ghana. Daily historical data for sunlight, humidity, rainfall, maximum and minimum temperature from 2000–2016 were obtained from the Ghana Meteorological Agency, Ghana. K-nearest neighbors (KNN) algorithm was used to compute the missing data points in the dataset. The yearly data points were calculated from the daily data of the selected weather variables using the arithmetic average. The yearly average temperature was computed from the arithmetic average of the yearly maximum and minimum temperature. Data for the weather variables was also taken from Bole, Tamale, and Yendi. These towns are part of the Northern Savanna and they are considered to be the food basket of Ghana. Due to the sensitivity of the ML algorithms used and the unequal weight of the data sets, the sample data sets were set into an identical scale–the min-max normalization scale in the interval $[0, 1]$. The sample data sets were divided into training (80%) and testing (20%) data sets. The training data set was used to build the classification ensemble algorithm and the testing data was used to validate the constructed ensemble model. The "accuracy"(accuracy is an evaluation metric used in classification and regression problems) of the training model is fined tuned using grid search. The training model was used to predict the crop yield signal (an increase crop yield or a decrease in crop yields). Using accuracy and the receiver operating characteristics curve (ROC) or area under the curve (AUC), the predictions on the testing data set were evaluated.

2.2. Model Evaluation and Feature Importance

Stacking is an ensemble learning technique that combines multiple classification or regression models through a meta-classifier/meta-regressor to a single predictive model to reduce bias (boosting), variance (bagging), and improve the accuracy of predictions (stacking). Using the training data set, the base level classifiers were trained. The meta-model was trained on the outputs of the base level algorithm as features. Stacking ensemble learning algorithms are heterogenous because the base level classifiers are made of different classification algorithms. In this study, Adapted boosting (AdaBoost) and artificial neural network (ANN), were used as the base classifier and gradient boosting machine (GBM) was used as the meta-classifier. Stacking ensemble algorithm is outlined in Algorithm 1.

Algorithm 1: Algorithm of Stacking Ensemble

    Input: $D : \{(x_i, y_i)\} \mid x_i \in X, y_i \in Y$

    Output: An ensemble classifier H

1 *Step 1. Learn base-level classifiers*

2 **for** $t \leftarrow 1$ to N **do**

3   Learn a base classifier h_t based on T, $h_t = H_t(T)$

4 **end for**

5 *Step 2. Construct new data set from T, $T' = \phi$*

6 **for** $i \leftarrow 1$ to n **do**

7   Construct a new data set that contains $\{x_i^{new}, y_i\}$
      where $\{x_i^{new}, y_i\} = \{h_t(x_i)$ for $t = 1$ to $N\}$

8 **end for**

9 *Step 3. Learn a second-level classifier*

10 Learn a new classifier h^{new} based on the newly constructed data set.

11 **Return** $H(x) = h^{new}(h_1(x), h_2(x), \cdots, h_N(x))$

The performance of the evaluation metrics of the proposed stacking ensemble classifier is presented in Table 1. The table gives the performance of the ensemble classifier on the testing data set. For binary classification, an AUC value of 50% or less is as good as randomly selecting the labels. An AUC value closer to 1 and an accuracy value closer to 100% indicate the superiority of the proposed model. In general, the classification model was very optimal in predicting the class of the target variable (crop yield) using the selected features (weather variables).

Table 1. Performance of evaluation metrics of the proposed stacking ensemble classification.

	Bole	Tamale	Yendi
Accuracy	0.8319	0.8207	0.8895
AUC	0.8104	0.7991	0.8334

From the feature importance in Table 2, average temperature, rainfall, and maximum temperature were the three most important features in the classification model used in predicting the yield of maize at Bole. In Tamale, the three most important variables that contributed to predicting the yield of maize were average temperature, minimum temperature, and rainfall. With importance values of 0.3017, 0.3006, and 0.2875, rainfall, average temperature, and maximum temperature were the three variables that had most effect in predicting maize yield at Yendi. Generally, from these results, average temperature can be considered as the most important underlying weather variable that affects the yield of maize in the Northern region of Ghana.

Table 2. Performance of evaluation metrics of the proposed stacking ensemble classification.

Feature	Bole		Tamale		Yendi	
	Importance Value	Rank	Importance Value	Rank	Importance Value	Rank
minT	0.0908	4	0.2001	2	0.2452	4
maxT	0.2017	3	0.1402	4	0.2875	3
aveT	0.3042	1	0.2075	1	0.3006	2
Rainfall	0.2552	2	0.1905	3	0.3017	1
Sunlight	0.0646	5	0.0625	6	0.2501	5
Humidity	0.0579	6	0.0983	5	0.2466	6

3. Temperature-Based Weather Derivatives

Even though the literature on weather derivatives has evolved rapidly over the past few decades, a consensus theoretical framework for evaluating weather derivatives has not been reached. This can be attributed to the fact that the underlying factors of WD are not tradable in the financial market and, hence, traditional pricing approaches cannot be used for the valuation of this product. Further, weather indices do not correlate strongly with the prices of other financial products in the financial market and, as a result, the underlying indices cannot be substituted for a linked exchange security.

3.1. Previous Temperature Dynamics Models

Alaton et al. [18] proposed the following mean-reverting model for the dynamics of temperature variations:

$$dT(t) = dS(t) + \beta(T(t) - S(t))dt + \sigma(t)dB(t), \tag{2}$$

where $T(t)$ represents the daily average temperature, $S(t)$ is the deterministic seasonal component, $B(t)$ is a Brownian motion, β is a constant mean-reversion rate, and $\sigma(t)$ is the volatility. They assumed that the volatility is a piecewise constant function which characterizes the monthly variation in volatility. From the proposed model of [18], Benth and Benth [20] suggested the following mean-reverting model for the time-dynamics of Norwegian daily average temperature:

$$dT(t) = dS(t) + \beta(T(t) - S(t))dt + \sigma(t)dL(t). \tag{3}$$

As indicated by [20], the only innovation from Equation (2) is the introduction of a Lévy noise, $L(t)$, instead of the Brownian motion $B(t)$. They used the generalized hyperbolic distribution, which is a special class of Lévy process, to capture the skewness and (semi-) heavy tails of their temperature data. However, to allow analytical pricing using the temperature dynamics model, Benth and Benth [21] later proposed to use a Brownian motion, instead of the Lévy process, in Equation (3). Clearly, References [18,20,21] used a constant mean-reversion rate, a volatility which is piecewise constant function, and they did not consider multi-dimensional locations in their models.

The introduction of a time-varying speed of mean-reversion, a local volatility that is able to capture the local variations of the daily average temperature at the selected locations, and a multi-dimensional daily average temperature model for multi-locations gave rise to the innovations in our proposed model and pricing. Our proposed models capture all the stylized facts of the selected locations and, to the best of our knowledge, it is the first of its kind in the literature.

3.2. Daily Average Temperature Data

The daily average temperature data (degree Celsius) for Bole and Tamale in the Northern savanna of Ghana, over a measurement period from 01/01/1992 to 31/08/2017, were taken from the Ghana Meteorological Service. The data consisted of the daily maximum and minimum temperatures. Missing data points were computed using KNN. The daily average temperature was computed from the arithmetic average of the daily minimum and maximum temperatures. For consistency in days (365 days) per year, February 29 was removed from the dataset for each leap year. As a result, the total daily average temperature had 9368 data points.

The seasonal and seasonally-adjusted (de-seasonalized) plot of the daily average temperature data for Bole and Tamale are presented in Figures 1 and 2, respectively. The spatial correlation between Bole and Tamale was estimated, in order to develop a basket temperature for weather derivatives. The de-seasonalized daily average temperature (as suggested by [24]) was used to estimate the correlation matrix between the two towns (Bole and Tamale). From Table 3, it is clear that the average temperature correlation between the selected locations was very high. This indicates that a maize farm at Bole will encounter the same weather-related risk as a maize farm in Tamale or Yendi, and vice versa. Hence, a weather station in any of these locations can be used for the contract without introducing

geographical basis risk in the contract. A farmer with two or more farms at these locations can buy a single temperature basket derivatives contract.
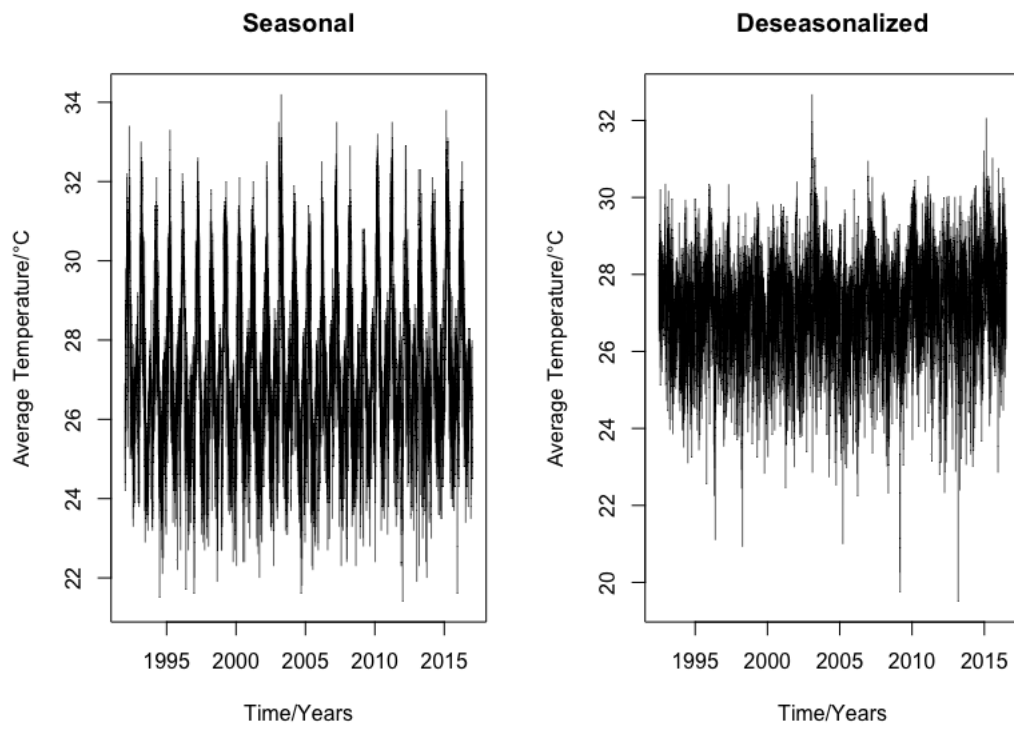

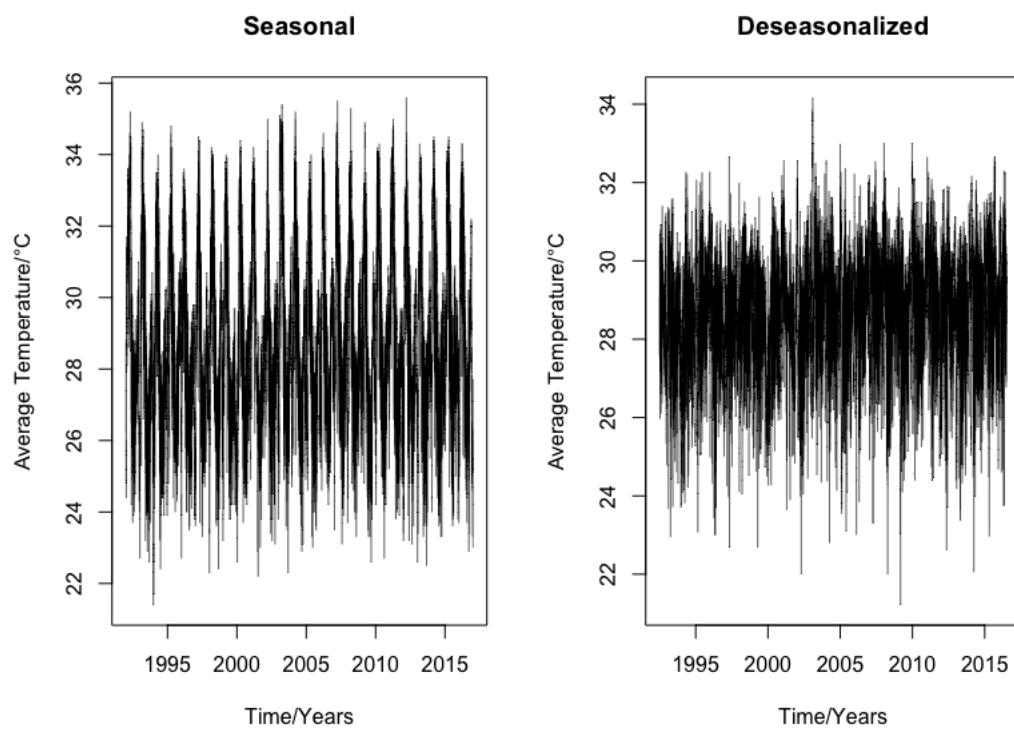**Figure 1.** Seasonal and de-seasonalized daily average temperature of Bole.

Figure 2. Seasonal and de-seasonalized daily average temperature of Tamale.

Table 3. Correlation matrix of the de-seasonalized daily average temperature.

	Bole	Tamale	Yendi
Bole	1	0.8733	0.8547
Tamale	0.8733	1	0.8998
Yendi	0.8547	0.8998	1

3.3. Stochastic Dynamics of Daily Average Temperature

Motivated by [18,20,25], we propose a new average temperature dynamics model which is able to capture major stylized facts of daily average temperature, such as locality features, seasonality, mean-reversion, and volatility. These stylized facts were consistent with the daily average temperature of the chosen location for this study. The proposed one-dimensional model was extended to a multi-dimensional temperature model to cover the selected locations under study. For convenience and analytical tractability, we assumed that the residuals of the daily average temperature were independently and identically distributed (i.i.d.) standard normal. The proposed daily average temperature model is given as

$$dT(t) = dS(t) + \beta(t)\big(T(t) - S(t)\big)dt + \sigma T(t)dB(t), \tag{4}$$

where $T(t)$ represents the daily average temperature, $S(t)$ is the deterministic seasonal component, $\beta(t)$ is the time-varying speed of mean-reversion, and $\sigma T(t)$ is the daily average temperature volatility through time.

Following [18], the seasonality component is defined as

$$S(t) = A + Bt + C \sin\left(\frac{2\pi t}{365} + \vartheta\right), \tag{5}$$

which is made up of a seasonal component $(C \sin 2\pi t/365 + \vartheta)$ and a trend component $(A + Bt)$, where A and B denote the constant and the coefficient of the linearity of the seasonal trend, respectively; C denotes the daily average temperature amplitude, ϑ; and t is the time, measured in days.

Equation (5) can be transformed to

$$S(t) = a + bt + c \, \sin\left(\frac{2\pi t}{365}\right) + d \, \cos\left(\frac{2\pi t}{365}\right). \tag{6}$$

By comparing (5) to (6), the relationship of the parameters is given below

$$A = a; \quad B = b; \quad C = \sqrt{c^2 + d^2}; \quad \vartheta = \arctan\left(\frac{d}{c}\right).$$

The numerical values of the constant in Equation (6) are estimated by fitting the function to the historical daily average temperature data using the method of least squares. The seasonal component for Bole and Tamale are given in the following function, respectively,

$$\begin{aligned} S(t) &= 22.15 + (4.57 \cdot 10^{-5})t + 1.98 \sin\left(\frac{2\pi t}{365} - 67.71\right), \\ S(t) &= 18.38 + (7.03 \cdot 10^{-5})t + 2.06 \sin\left(\frac{2\pi t}{365} - 72.89\right). \end{aligned} \tag{7}$$

Using additive seasonal decomposition by moving averages, the daily average temperature was decomposed into a seasonal, linear, and a random (residual) component as shown in Figures 3–5 respectively. The trend component (Figure 4) was smoother than the actual daily average temperature data plot (Figures 1 and 2, seasonalized) and captured the main movement of the daily average

temperature data without the minor variations. The estimated seasonal component of the daily average temperature is presented in Figure 6.
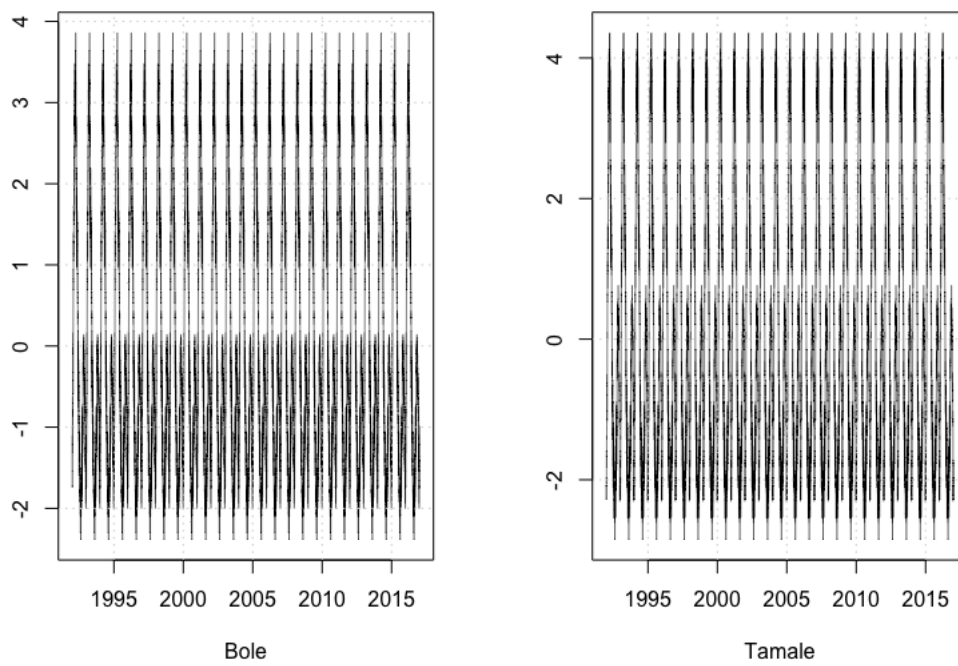

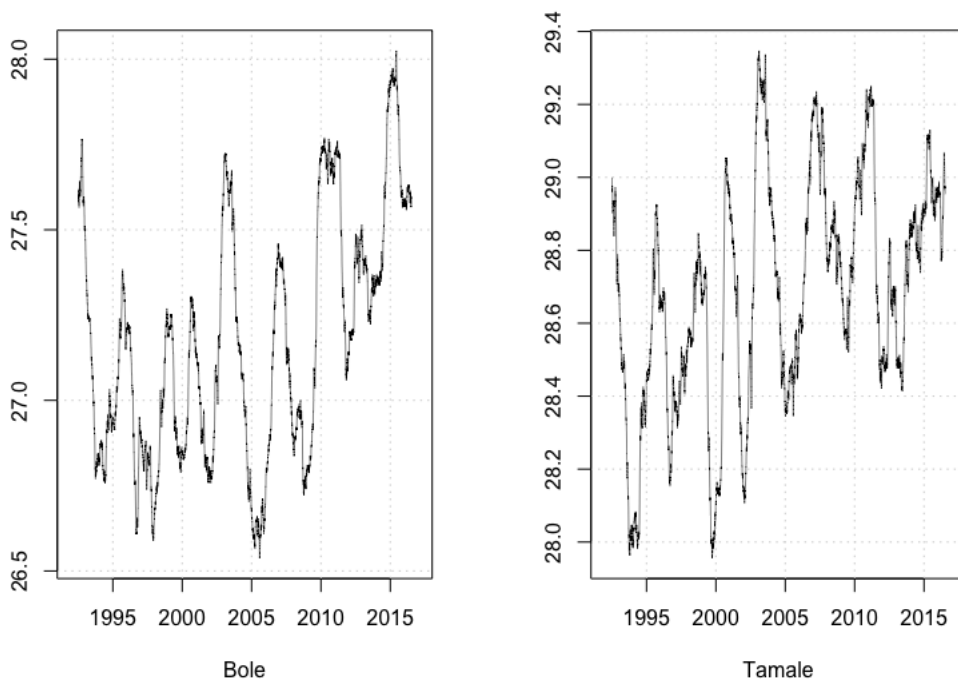**Figure 3.** Seasonal trend of the daily average temperature for Bole and Tamale.

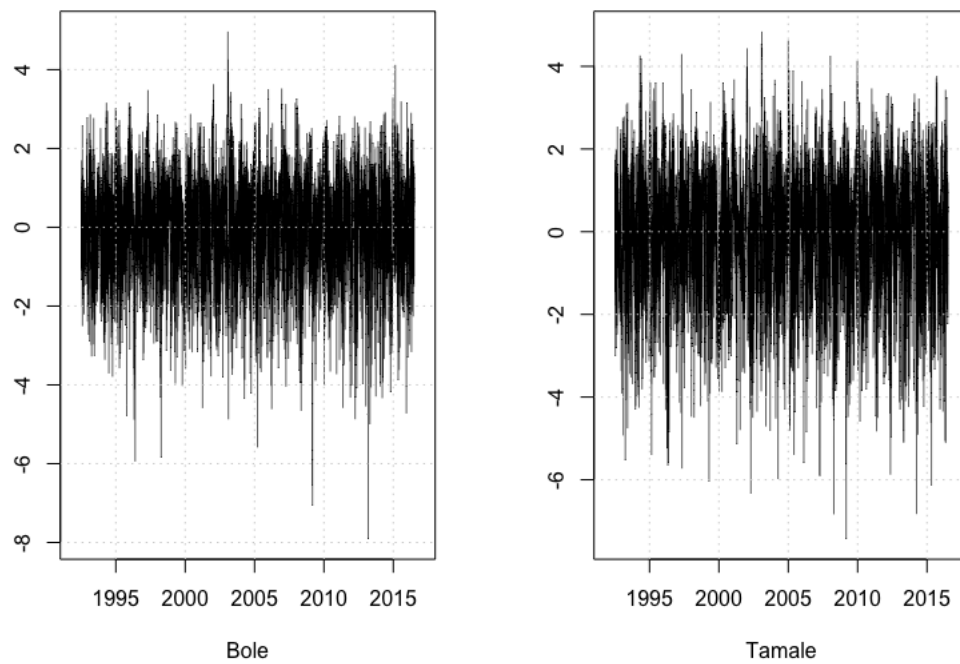**Figure 4.** Linear trend of the daily average temperature for Bole and Tamale.

Figure 5. Residuals of the daily average temperature for Bole and Tamale.
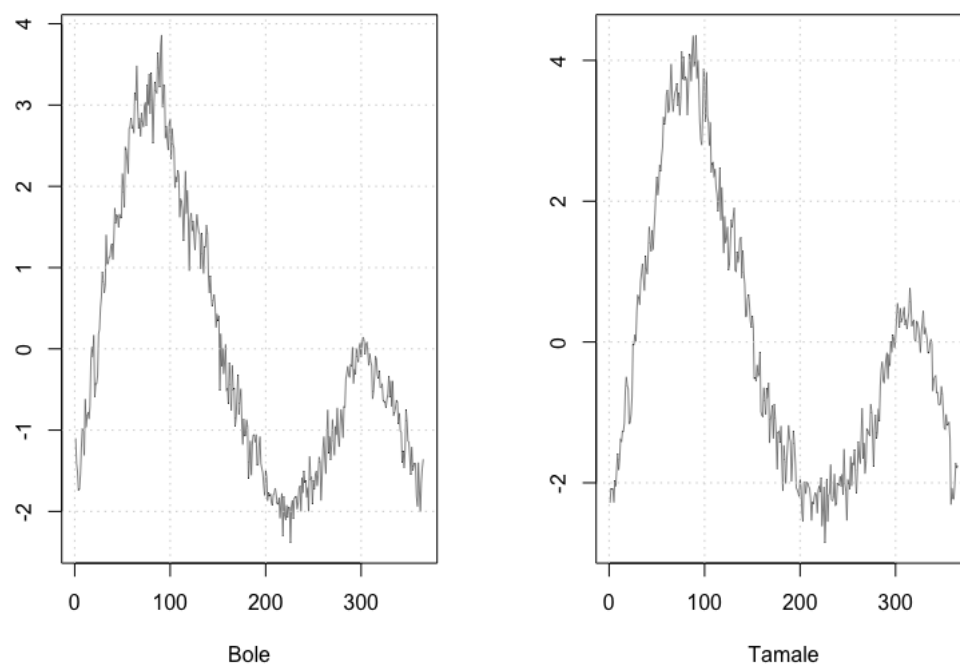

Figure 6. Estimated seasonal figure of the daily average temperature for Bole and Tamale.

3.4. Temperature-Based Weather Derivative Pricing

Theorem 1 (Girsanov Theorem). *Let B_t be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\lambda = \{\lambda_t : 0 \leq t \leq T\}$ be an adaptive process satisfying the Novikov condition*

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^t \lambda_u^2 du\right)\right] < \infty. \tag{8}$$

Let　　$Z(t) = \exp\left(\int_0^t \lambda_u dB_u - \frac{1}{2}\int_0^t \lambda_u^2 du\right).$

Then, $\mathbb{Q} \sim \mathbb{P}$ can be determined by the Radon–Nikodym derivative

$$\frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathcal{F}_t = Z(T). \tag{9}$$

Then, the random process

$$W_t = B_t - \int_0^t \lambda_s ds,$$
$$dW_t = dB_t - \lambda_t dt, \tag{10}$$

is a standard Brownian motion under the measure $\mathbb{Q}^\lambda$.

Remark 1. *The Novikov condition in the Girsanov theorem ensures that Z is positive martingale whenever $\mathbb{E}(Z) = 1$. This is referred to as the Radon–Nikodym derivative.*

Remark 2. *λ is refered to as the market price of risk (MPR). As there is no real weather derivative market in Africa from which the prices can be obtained, λ is assumed to be a constant. For a constant λ, Equation (10) can be re-defined as*

$$dW_t = dB_t - \lambda dt. \tag{11}$$

Lemma 1. *If the daily average temperature follows the proposed model in Equation (4), then the explicit solution is given by*

$$T_t = S_t + (T_0 - S_0)e^{\int_0^t \beta_s ds} + e^{\int_0^t \beta_s ds} \int_0^t \sigma T_u e^{\int_0^t \beta_s ds} dB_u. \tag{12}$$

Proof. We have that

$$dT_t = dS_t + \beta_t(T_t - S_t)dt + \sigma T_t dB_t$$
$$d\tilde{T}_t = \beta_t \tilde{T}_t + \sigma T_t dB_t, \tag{13}$$

where $\tilde{T} = T_t - S_t$. Using the transformation below, $d\tilde{T}_t$ can be evaluated,

$$F(\tilde{T}_t, t) = \tilde{T}_t e^{-\int_0^t \beta_s ds}$$

$$\frac{\partial F}{\partial \tilde{T}_t} = e^{-\int_0^t \beta_s ds}; \qquad \frac{\partial^2 F}{\partial \tilde{T}_t^2} = 0; \qquad \frac{\partial F}{\partial t} = -\beta_t \tilde{T}_t e^{-\int_0^t \beta_s ds}.$$

Applying Itô's Lemma and Equation (13),

$$dF_t = \sigma T_t e^{-\int_0^t \beta_s ds} dB_t. \tag{14}$$

Integrating Equation (14) over the interval $[0, t]$,

$$F_t = F_0 + \int_0^t \sigma T_u e^{-\int_0^u \beta_s ds} dB_u$$

$$\tilde{T}_t e^{-\int_0^t \beta_s ds} = \tilde{T}_0 + \int_0^t \sigma T_u e^{-\int_0^u \beta_s ds} dB_u$$

$$\tilde{T}_t = \tilde{T}_0 e^{\int_0^t \beta_s ds} + e^{\int_0^t \beta_s ds} \int_0^t \sigma T_u e^{-\int_0^u \beta_s ds} dB_u$$

$$T_t = S_t + (T_0 - S_0)e^{\int_0^t \beta_s ds} + e^{\int_0^t \beta_s ds} \int_0^t \sigma T_u e^{-\int_0^u \beta_s ds} dB_u.$$

□

Lemma 2. *Under the risk-neutral measure $\mathbb{Q}$, the explicit solution of the daily average temperature model is given by*

$$T_t = S_t + (T_0 - S_0)e^{\int_0^t \beta_s ds} + \int_0^t \sigma \lambda T_u e^{\int_u^t \beta_s ds} du + \int_0^t \sigma T_u e^{\int_u^t \beta_s ds} dW_u. \tag{15}$$

Proof. By substituting Equation (11) into Equation (13) and following the steps of the proof of Lemma 12, the lemma can be derived. □

3.4.1. CAT Futures and Options on Futures

Suppose that, for a contract period $[t_1, t_2]$, the temperature dynamics follow the TML model. Then, there is a price dynamic of futures written on a CAT index with $t \leq t_1 < t_2$. The futures price of CAT is given by

$$0 = e^{-r(t_2-t)} \mathbb{E}_{\mathbb{Q}} \left[\int_{t_1}^{t_2} T_x dx - F_{CAT}(t, t_1, t_2) \mid \mathcal{F}_t \right]. \tag{16}$$

As the future price $F(t, t_1, t_2)$ is $\mathcal{F}_t$ adapted under the measure $\mathbb{Q}$,

$$F_{CAT}(t, t_1, t_2) = \mathbb{E}_{\mathbb{Q}} \left[\int_{t_1}^{t_2} T_x dx \mid \mathcal{F}_t \right]. \tag{17}$$

Proposition 1. *Suppose the daily average temperature follows Model (4). Then, the price of CAT-futures at time $t \leq t_1 \leq t_2$ for the contract period $[t_1, t_2]$ is given by:*

$$F_{CAT}(t, t_1, t_2) = \int_{t_1}^{t_2} S_x dx + \int_{t_1}^{t_2} (T_t - S_t)e^{\int_t^x \beta_s ds} dx + L_1,$$

where $L_1 = \int_t^{t_1} \int_{t_1}^{t_2} e^{\int_0^x \beta_s ds} \sigma \lambda T_u e^{\int_u^0 \beta_s ds} dx du + \int_{t_2}^{t_1} \int_{t_1}^{t_2} e^{\int_0^x \beta_s ds} \sigma \lambda T_u e^{\int_u^0 \beta_s ds} dx du.$

Proof. From Equation (17),

$$F_{CAT}(t, t_1, t_2) = \mathbb{E}_{\mathbb{Q}} \left[\int_{t_1}^{t_2} T_x dx \mid \mathcal{F}_t \right].$$

From Lemma 2 and for any time $x \geq t$

$$T_x = S_x + (T_t - S_t)e^{\int_t^x \beta_s ds} + \int_t^x \sigma \lambda T_u e^{\int_u^x \beta_s ds} du + \int_t^x \sigma T_u e^{\int_u^x \beta_s ds} dW_u$$

$$F_{CAT}(t, t_1, t_2) = \mathbb{E}_{\mathbb{Q}} \left[\int_{t_1}^{t_2} \left(S_x + (T_t - S_t)e^{\int_t^x \beta_s ds} + \int_t^x \sigma \lambda T_u e^{\int_u^x \beta_s ds} du + \int_t^x \sigma T_u e^{\int_u^x \beta_s ds} dW_u \right) dx \mid \mathcal{F}_t \right]$$

$$= \int_{t_1}^{t_2} S_x dx + \int_{t_1}^{t_2} (T_t - S_t)e^{\int_t^x \beta_s ds} dx + \int_{t_1}^{t_2} \int_t^x \sigma \lambda T_u e^{\int_u^x \beta_s ds} du dx$$

$$= \int_{t_1}^{t_2} S_x dx + \int_{t_1}^{t_2} (T_t - S_t)e^{\int_t^x \beta_s ds} dx + L_1,$$

where

$$L_1 = \int_{t_1}^{t_2} \int_t^x \sigma \lambda T_u e^{\int_u^x \beta_s ds} du dx$$

$$= \int_{t_1}^{t_2} \int_t^{t_2} \mathbb{1}_{[t,x]}(u) \sigma \lambda T_u e^{\int_u^x \beta_s ds} du dx$$

$$= \int_t^{t_2} \int_{t_1}^{t_2} \mathbb{1}_{[t,x]}(u) \sigma \lambda T_u e^{\int_u^x \beta_s ds} dx du$$

$$= \int_t^{t_1} \int_{t_1}^{t_2} \mathbb{1}_{[t,x]}(u) \sigma \lambda T_u e^{\int_u^x \beta_s ds} dx du + \int_{t_1}^{t_2} \int_{t_1}^{t_2} \mathbb{1}_{[t,x]}(u) \sigma \lambda T_u e^{\int_u^x \beta_s ds} dx du$$

$$= \int_t^{t_1} \int_{t_1}^{t_2} e^{\int_0^x \beta_s ds} \sigma \lambda T_u e^{\int_u^0 \beta_s ds} dx du + \int_u^{t_2} \int_{t_1}^{t_2} e^{\int_0^x \beta_s ds} \sigma \lambda T_u e^{\int_u^0 \beta_s ds} dx du.$$

$\square$

Proposition 2. *At time $t_1 \leq t \leq t_2$, the in-period (in the contract) valuation of the CAT futures is given by*

$$F_{CAT}(t, t_1, t_2) = \int_{t_1}^{t} T_x dx + \int_{t}^{t_2} S_x dx + \int_{t}^{t_2} (T_t - S_t) e^{\int_{t}^{x} \beta_s ds} dx + \int_{u}^{t_2} \int_{t}^{t_2} e^{\int_{0}^{x} \beta_s ds} \sigma \lambda T_u e^{\int_{u}^{0} \beta_s ds} dx du.$$

Proof. From the CAT futures price in Equation (17),

$$\begin{aligned} F_{CAT}(t, t_1, t_2) &= \mathbb{E}_{\mathbb{Q}} \left[\int_{t_1}^{t_2} T_x dx \mid \mathcal{F}_t \right] \\ &= \mathbb{E}_{\mathbb{Q}} \left[\left(\int_{t_1}^{t} T_x dx + \int_{t}^{t_2} T_x dx \right) \Big| \mathcal{F}_t \right] \\ &= \int_{t_1}^{t} T_x dx + \mathbb{E}_{\mathbb{Q}} \left[\int_{t}^{t_2} T_x dx \mid \mathcal{F}_t \right] \\ &= \int_{t_1}^{t} T_x dx + F_{CAT}(t, t, t_2). \end{aligned}$$

From Proposition 1,

$$F_{CAT}(t, t_1, t_2) = \int_{t_1}^{t} T_x dx + \int_{t}^{t_2} S_x dx + \int_{t}^{t_2} (T_t - S_t) e^{\int_{t}^{x} \beta_s ds} dx + \int_{u}^{t_2} \int_{t}^{t_2} e^{\int_{0}^{x} \beta_s ds} \sigma \lambda T_u e^{\int_{u}^{0} \beta_s ds} dx du.$$

$\square$

Lemma 3. *The dynamics of the CAT futures under the equivalent probability measure $\mathbb{Q}$ and measured over the contract period $[t_1, t_2]$ is given by*

$$dF_{CAT}(t, t_1, t_2) = \Sigma_{CAT}(t, t_1, t_2, T_t) dW_t,$$

where

$$\Sigma_{CAT}(t, t_1, t_2, T_t) = \sigma T_t \int_{t_1}^{t_2} e^{\int_{t}^{x} \beta_s ds} dx.$$

Proof. We have that

$$\begin{aligned} \frac{dF_{CAT}(t, t_1, t_2)}{dT_t} &= \int_{t_1}^{t_2} e^{\int_{t}^{x} \beta_s ds} dx \\ dF_{CAT}(t, t_1, t_2) &= \int_{t_1}^{t_2} e^{\int_{t}^{x} \beta_s ds} dx dT_t \\ dF_{CAT}(t, t_1, t_2) &= \sigma T_t \int_{t_1}^{t_2} e^{\int_{t}^{x} \beta_s ds} dx dW_t. \end{aligned}$$

$\square$

Proposition 3. *The call option price at exercise time t_n and strike price C is given by*

$$C_{CAT}(t, t_n, t_1, t_2) = e^{-r(t_n - t)} \left((F_{CAT}(t, t_1, t_2) - C) \Phi(\Delta(t, t_n, t_1, t_2, T_t)) + \Sigma_{t, t_n} \phi(\Delta(t, t_n, t_1, t_2, T_t)) \right),$$

where

$$\Delta(t, t_n, t_1, t_2, T_t) = \frac{F_{CAT}(t, t_1, t_2) - C}{\sqrt{\Sigma_{t, t_n}^2}}; \qquad \Sigma_{t, t_n}^2 = \int_{t}^{t_n} \Sigma_{CAT}^2(s, t_1, t_2, T_t) ds,$$

Φ is the cumulative standard normal distribution function, ϕ is the standard normal density function, and $\phi(\cdot) = \Phi'(\cdot)$.

Proof. By definition, a call option price at exercise time t_n and strike price C is given as

$$C_{CAT}(t, t_n, t_1, t_2) = e^{-r(t_n-t)}\mathbb{E}_\mathbb{Q}\big[\max\big(F_{CAT}(t_n, t_1, t_2) - C, 0\big)\big|\mathcal{F}_t\big].$$

From Lemma 3,

$$\int_t^{t_n} dF_{CAT}(s, t_1, t_2) = \int_t^{t_n} \Sigma_{CAT}(s, t_1, t_2, T_t)dW_s,$$

$$F_{CAT}(t_n, t_1, t_2) = F_{CAT}(t, t_1, t_2) + \int_t^{t_n} \Sigma_{CAT}(s, t_1, t_2, T_t)dW_s.$$

$F_{CAT}(t_n, t_1, t_2)$ is normally distributed under $\mathbb{Q}^\lambda$, with expectation

$$\mathbb{E}_\lambda[F_{CAT}(t_n, t_1, t_2)] = F_{CAT}(t, t_1, t_2)$$

and variance

$$Var_\lambda[F_{CAT}(t_n, t_1, t_2)] = \int_t^{t_n} \Sigma_{CAT}^2(s, t_1, t_2, T_t)ds = \Sigma_{t,t_n}.$$

Hence,

$$\begin{aligned}
C_{CAT}(t, t_n, t_1, t_2) &= e^{-r(t_n-t)}\mathbb{E}_\mathbb{Q}\big[\max\big(F_{CAT}(t_n, t_1, t_2) - C, 0\big)\big|\mathcal{F}_t\big]\\
&= e^{-r(t_n-t)}\int_C^\infty (y-C)f_{CAT}(y)dy\\
&= e^{-r(t_n-t)}\Big(\big(F_{CAT}(t, t_1, t_2) - C\big)\Phi(\Delta(t, t_n, t_1, t_2, T_t)) + \Sigma_{t,t_n}\phi(\Delta(t, t_n, t_1, t_2, T_t))\Big),
\end{aligned}$$

where $\Delta(t, t_n, t_1, t_2, T_t) = \dfrac{F_{CAT}(t, t_1, t_2) - C}{\sqrt{\Sigma_{t,t_n}^2}}$. $\quad\square$

3.4.2. GDD Futures and Options on Futures

Similar to the definition of the CAT future price, the GDD future price is given as

$$0 = e^{-r(t_2-t)}\mathbb{E}_\mathbb{Q}\left[\int_{t_1}^{t_2} T_x dx - F_{GDD}(t, t_1, t_2)\bigg|\mathcal{F}_t\right].$$

Using the same idea in deriving the CAT futures price, the price of the GDD-futures can be derived as

$$F_{GDD}(t, t_1, t_2) = \mathbb{E}_\mathbb{Q}\left[\int_{t_1}^{t_2} \max\big(T_x - T^{optimal}, 0\big)dx\bigg|\mathcal{F}_t\right], \tag{18}$$

where $T^{optimal}$ is the optimal normal temperature at which a crop will develop.

Proposition 4. *Suppose the daily average temperature follows Model (4). Then, the price of GDD-futures at time $t \le t_1 \le t_2$ for the contract period $[t_1, t_2]$ is given by:*

$$WehavethatF_{GDD}(t, t_1, t_2) = \int_{t_1}^{t_2} \Psi(t, x)\Big[\phi\big(\Delta(t, x)\big) + \Delta(t, x)\Phi\big(\Delta(t, x)\big),$$

where

$$\Psi^2(t,x) = \int_t^x \sigma^2 T_u^2 e^{2\int_u^x \beta_s ds} du; \qquad \Delta(t,x) = \frac{S_x + (T_t - S_t)e^{\int_t^x \beta_s ds} + \int_t^x \sigma\lambda T_u e^{\int_u^x \beta_s ds} du - C}{\Psi(t,x)}.$$

Proof. By definition.

$$GDD = \int_{t_1}^{t_2} \max(T_x - C, 0)dx,$$

$$F_{GDD}(t,t_1,t_2) = \mathbb{E}_{\mathbb{Q}}\left[\int_{t_1}^{t_2} \max\left(T_x - C, 0\right)dx \,\Big|\, \mathcal{F}_t \right].$$

(19)

Recall, from Lemma 2 and for any time $x \geq t$,

$$T_x = S_x + (T_t - S_t)e^{\int_t^x \beta_s ds} + \int_t^x \sigma\lambda T_u e^{\int_u^x \beta_s ds}du + \int_t^x \sigma T_u e^{\int_u^x \beta_s ds}dW_u,$$

(20)

$$T_x = D_x = A(t,x) + B(t,x),$$

(21)

where

$$D(t,x) = S_x + (T_t - S_t)e^{\int_t^x \beta_s ds} + \int_t^x \sigma\lambda T_u e^{\int_u^x \beta_s ds}du, \text{ and } B(t,x) = \int_t^x \sigma T_u e^{\int_u^x \beta_s ds}dW_u.$$

The distribution of D_x can be determined:
$A(t,x)$ is deterministic and, hence,

$$B(t,x) \sim N\left(0, \int_t^x \sigma^2 T_u^2 e^{2\int_u^x \beta_s ds}du\right) = N\left(0, \Psi^2(t,x)\right).$$

It follows, from Equation (21), that

$$D_x \sim N\left(A(t,x), \Psi^2(t,x)\right).$$

Thus, D_x can be written, in terms of the standard normal variable $Z \sim N(0,1)$, as

$$D(t,x) = A(t,x) + \left(\Psi^2(t,x)\right)^{\frac{1}{2}}Z.$$

(22)

Consider

$$T_x - C > 0.$$

This requires

$$\left(\Psi^2(t,x)\right)^{\frac{1}{2}}Z > C - A(t,x),$$

$$Z > \frac{C - A(t,x)}{\left(\Psi^2(t,x)\right)^{\frac{1}{2}}Z} = \Delta_1(t,x).$$

(23)

From Equation (23),

$$C = A(t,x) + \Delta(t,x)\left(\Psi^2(t,x)\right)^{\frac{1}{2}}Z.$$

(24)

From Equations (19) and (23),

$$\mathbb{E}_{\mathbb{Q}}\left[\int_{t_1}^{t_2} \max\left(T_x - C, 0\right)dx \,\Big|\, \mathcal{F}_t \right] = \int_{\Delta(t,x)}^{+\infty} \left(D(t,x) - C\right)\frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}}dz.$$

(25)

Substituting Equations (22) and (24) into Equation (25),

$$
= \int_{\Delta(t,x)}^{+\infty} \left(A(t,x) + \left(\Psi^2(t,x)\right)^{\frac{1}{2}} Z - A(t,x) - \Delta(t,x)\left(\Psi^2(t,x)\right)^{\frac{1}{2}} Z \right) \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}} dz
$$

$$
= \int_{\Delta(t,x)}^{+\infty} \left(\left(\Psi^2(t,x)\right)^{\frac{1}{2}} z - \Delta(t,x)\left(\Psi^2(t,x)\right)^{\frac{1}{2}} z \right) \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}} dz
$$

$$
= \left(\Psi^2(t,x)\right)^{\frac{1}{2}} \left(\int_{\Delta_1(t,x)}^{+\infty} \frac{z e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}} dz + \Delta_1(t,x)\Phi(-\Delta_1(t,x)) \right)
$$

$$
= \left(\Psi^2(t,x)\right)^{\frac{1}{2}} \left(\frac{e^{-\frac{1}{2}\Delta(t,x)^2}}{\sqrt{2\pi}} dz + \Delta(t,x)\Phi(\Delta(t,x)) \right)
$$

$$
= \left(\Psi^2(t,x)\right)^{\frac{1}{2}} \left[\phi(\Delta(t,x)) + \Delta(t,x)\Phi(\Delta(t,x)) \right],,
$$

where

$$
\Delta(t,x) = -\Delta_1(t,x) = \frac{G(t,x) - C}{\left(\Psi^2(t,x)\right)^{\frac{1}{2}}},
$$

$$
\mathbb{E}_{\mathbb{Q}}\left[\int_{t_1}^{t_2} \max\left(T_x - C, 0\right) dx \middle| \mathcal{F}_t \right] = \Psi(t,x)\left[\phi\big(\Delta(t,x)\big) + \Delta(t,x)\Phi\big(\Delta(t,x)\big) \right]. \tag{26}
$$

Substituting Equation (26) into Equation (19) gives the Proposition. □

Lemma 4. *The dynamics of the GDD futures under the equivalent probability measure $\mathbb{Q}$ measured over the period $[t_1, t_2]$ are given by*

$$
dF_{GDD}(t, t_1, t_2) = \Pi_{GDD}(t, t_1, t_2)dW_t,
$$

where Π_{GDD} is called the term structure of the GDD futures volatility,

$$
\Pi_{GDD} = \sigma T_t \int_{t_1}^{t_2} e^{\int_t^x \beta_s ds} \Phi\left(\frac{h(t, x, e^{\int_t^x \beta_s ds}(T_t - S_t))}{\Psi(t,x)} \right) ds; \quad \Psi^2(t,x) = \int_t^x \sigma^2 T_u^2 e^{2\int_u^x \beta_s ds} du;
$$

$$
h(t, x, e^{\int_t^x \beta_s ds}(T_t - S_t)) = S_x + (T_t - S_t)e^{\int_t^x \beta_s ds} + \int_t^x \sigma \lambda T_u e^{\int_u^x \beta_s ds} du - C.
$$

Proof. Let $h(t, x, e^{\int_t^x \beta_s ds}(T_t - S_t)) = S_x + (T_t - S_t)e^{\int_t^x \beta_s ds} + \int_t^x \sigma \lambda T_u e^{\int_u^x \beta_s ds} du - C$

$$
\frac{dh}{dT_t} = e^{\int_t^x \beta_s ds}
$$

From Proposition 4,

$$
\frac{dF_{GDD}}{dT_t} = \int_{t_1}^{t_2} \Psi(t,x)\Upsilon'\left(\frac{h(t, x, e^{\int_t^x \beta_s ds}(T_t - S_t))}{\Psi(t,x)} \right) \frac{h'(t, x, e^{\int_t^x \beta_s ds}(T_t - S_t))}{\Psi^2(t,x)} ds
$$

$$
= \int_{t_1}^{t_2} \Upsilon'\left(\frac{h(t, x, e^{\int_t^x \beta_s ds}(T_t - S_t))}{\Psi(t,x)} \right) h'(t, x, e^{\int_t^x \beta_s ds}(T_t - S_t)) ds
$$

$$
= \int_{t_1}^{t_2} e^{\int_t^x \beta_s ds} \Phi\left(\frac{h(t, x, e^{\int_t^x \beta_s ds}(T_t - S_t))}{\Psi(t,x)} \right) ds
$$

$$
dF_{GDD} = \sigma T_t \int_{t_1}^{t_2} e^{\int_t^x \beta_s ds} \Phi\left(\frac{h(t, x, e^{\int_t^x \beta_s ds}(T_t - S_t))}{\Psi(t,x)} \right) ds dW_t
$$

$$
dF_{GDD} = \Pi_{GDD}(t, t_1, t_2)dW_t,
$$

where

$$Y(\Delta(t,x)) = \phi(\Delta(t,x)) + \Delta(t,x)\Phi(\Delta(t,x)); \quad \Delta(t,x) = \frac{h(t,x,e^{\int_t^x \beta_s ds}(T_t - S_t))}{\Psi(t,x)};$$

$$\Pi_{GDD} = \sigma T_t \int_{t_1}^{t_2} e^{\int_t^x \beta_s ds} \Phi\left(\frac{h(t,x,e^{\int_t^x \beta_s ds}(T_t - S_t))}{\Psi(t,x)}\right) ds.$$

$\square$

Proposition 5. *For a strike price C and maturity time $t \le t_n \le t_1$, the price of a call option at time t on a GDD futures contract is given by*

$$C_{GDD}(t,t_n,t_1,t_2) = e^{-r(t_n-t)}\mathbb{E}_{\mathbb{Q}}\left[\max\left(\int_{t_1}^{t_2}\Psi(t,x)P(t,x,t_n,(T_t-S_t))ds - C, 0\right)\right],$$

where

$$P(t,x,t_n,(T_t-S_t)) = \tilde{Y}\left(t,x,e^{\int_t^x \beta_s ds}(T_t - S_t) + \int_t^{t_n}\lambda\sigma e^{\int_u^x \beta_s ds}du + \Sigma(x,t,t_n)Y\right)$$

$$\tilde{Y}(t,x,e^{\int_t^x \beta_s ds}(T_t - S_t)) = Y\left(\frac{h(t,x,e^{\int_t^x \beta_s ds}(T_t - S_t))}{\Psi(t,x)}\right)$$

and

$$\Sigma(x,t,t_n) = \int_t^{t_n}\sigma^2 T_u^2 e^{\int_u^x \beta_s ds}du.$$

Proof. By definition ,

$$C_{GDD}(t,t_n,t_1,t_2) = e^{-r(t_n-t)}\mathbb{E}_{\mathbb{Q}}\left[\int_{t_1}^{t_2}\max\left(F_{GDD}(t_n,t_1,t_2) - C, 0\right)dx\Big|\mathcal{F}_t\right]$$

$$F_{GDD}(t_n,t_1,t_2) = \int_{t_1}^{t_2}\Psi(t,x)\tilde{Y}\left(t,x,e^{\int_{t_n}^x \beta_s ds}(T_{t_n} - S_{t_n})\right)ds,$$

$$= \int_{t_1}^{t_2}\Psi(t,x)\tilde{Y}\left(t,x,e^{\int_{t_n}^x \beta_s ds}(T_t - S_t) + \int_t^{t_n}\lambda\sigma T_u e^{\int_u^x \beta_s ds}du + \int_t^{t_n}\sigma T_u e^{\int_u^x \beta_s ds}dW_u\right)ds$$

$$C_{GDD}(t,t_n,t_1,t_2) = e^{-r(t_n-t)}\mathbb{E}_{\mathbb{Q}}\left[\int_{t_1}^{t_2}\max\left(\int_{t_1}^{t_2}\Psi(t,x)\tilde{Y}\left(t,x,e^{\int_{t_n}^x \beta_s ds}(T_t - S_t) + \int_t^{t_n}\lambda\sigma T_u e^{\int_u^x \beta_s ds}du + \right.\right.\right.$$

$$\left.\left.\left.\int_t^{t_n}\sigma T_u e^{\int_u^x \beta_s ds}dW_u\right)ds - C, 0\right)dx\Big|\mathcal{F}_t\right],$$

where

$$\tilde{Y}(t,x,e^{\int_t^x \beta_s ds}(T_t - S_t)) = Y\left(\frac{h(t,x,e^{\int_t^x \beta_s ds}(T_t - S_t))}{\Psi(t,x)}\right).$$

$\square$

3.4.3. CAT and GDD Futures on Temperature Basket

Assuming N is the spatial locations in the basket, then $(\omega_i)_{i=1}^N$ will be the collection of weights for the spatial locations $(y_i)_{i=1}^N$. The basket of the daily average temperature at the N spatial locations for a given time t is defined as:

$$M(t) := \sum_{i=1}^N \omega_i T_i(t), \tag{27}$$

where $\sum_{i=1}^N \omega_i = 1$.

Assume the daily average temperature is spatially correlated across the random noise term and the risk-neutral distribution of the daily average temperature for the spatial locations is normally distributed in Model (4). Hence, the weighted sum of a normally distributed basket is also normally distributed. From the above assumptions, a new daily average temperature model for each spatial location y^i can be proposed,

$$dT_i(t) = dS_i(t) + \beta_i(t)\big(T_i(t) - S_i(t)\big)dt + \sigma T_i(t)dB_i(t). \tag{28}$$

Expressing Equation (28), for locations $i = 1, 2, 3 \cdots, N$, as an N-dimensional system leads to the equation below,

$$d\boldsymbol{T}(t) = d\boldsymbol{S}(t) + \boldsymbol{\beta}(t)\big(\boldsymbol{T}(t) - \boldsymbol{S}(t)\big)dt + \sigma\boldsymbol{T}(t)d\boldsymbol{B}(t), \tag{29}$$

where $\boldsymbol{B}(t) \sim N(0, \boldsymbol{\Omega}t)$ and $\boldsymbol{\Omega}$ is a covariance matrix. Using the linear transformation of multivariate normal distributions property,

$$Y \sim N(\mu, \Sigma) \Rightarrow DY \sim N(D\mu, D\Sigma D^T).$$

Suppose $Z \sim N(0, \boldsymbol{I}t)$ and $Y = DZ$. Then, it follows that $Y \sim N(0, DD^T t)$. Applying Cholesky factorization to Σ, we can derive a lower triangular form for D, and $\boldsymbol{W}_t$ will be expressed as an N-dimensional Brownian motion $\boldsymbol{V}_t$,

$$\boldsymbol{B}(t) = \boldsymbol{L}\boldsymbol{V}(t). \tag{30}$$

Then, $\boldsymbol{L}\boldsymbol{L}^T = \boldsymbol{\Omega}$, $\boldsymbol{L}$ is a lower triangular matrix with non-negative diagonal entries, $\boldsymbol{L}^T$ is an upper triangular matrix, and $\boldsymbol{V}(t) = (V_1(t), V_2(t), V_3(t), \cdots, V_N(t))^T$ with $dV_i(t)dV_j(t) = \delta_{ij}dt$. Equation (29) can be reformulated, in terms of $\boldsymbol{V}(t)$, as

$$d\boldsymbol{T}(t) = d\boldsymbol{S}(t) + \boldsymbol{\beta}(t)\big(\boldsymbol{T}(t) - \boldsymbol{S}(t)\big)dt + \sigma\boldsymbol{T}(t)\boldsymbol{L}d\boldsymbol{V}(t). \tag{31}$$

3.4.4. Girsanov's Theorem in $\mathbb{R}^N$

Let $\boldsymbol{V}(t) = (V_1(t), V_2(t), V_3(t), \cdots, V_N(t))$ be an N-dimensional Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\boldsymbol{\lambda} = (\lambda_1(t), \lambda_2(t), \lambda_3(t), \cdots, \lambda_N(t))$ be an N-dimensional adapted process on $[0, T]$.

Define

$$Z_{\boldsymbol{\lambda}}(t) := \exp\left(\int_0^t \boldsymbol{\lambda}(s)d\boldsymbol{V}(s) - \frac{1}{2}\int_0^t ||\boldsymbol{\lambda}(s)||^2 \, ds\right), \tag{32}$$

where $||\boldsymbol{\lambda}(s)||^2 = \sum_{i=1}^N \lambda_i(s)^2$.

Let

$$\tilde{\boldsymbol{V}}(t) = \boldsymbol{V}(t) + \int_0^t \boldsymbol{\lambda}(s)ds. \tag{33}$$

The component process of $\tilde{\boldsymbol{B}}(t)$ is independent under the measure $\mathbb{Q}$.

Suppose that

$$\mathbb{E}\int_0^t ||\lambda(s)||^2 \, Z(s)^2 ds < \infty.$$

Then, $\tilde{\boldsymbol{B}}(t)$ is an N-dimensional standard Brownian motion under the measure $\mathbb{Q}$, defined as

$$\frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|\mathcal{F}_t = Z(T). \tag{34}$$

Let

$$Z_{\boldsymbol{\lambda}}(t) := \exp\left(\int_0^t (\sigma\boldsymbol{T}(t)\boldsymbol{L})\boldsymbol{\lambda}(s)d\boldsymbol{V}(s) - \frac{1}{2}\int_0^t ||\sigma\boldsymbol{T}(s)\boldsymbol{L}||^{-2}||\boldsymbol{\lambda}(s)||^2 \, ds\right). \tag{35}$$

For a constant market price of risk at each geographical reference location in equation (35), it can be deduced that

$$W_\lambda(t) = V(t) - \int_0^t (\sigma T(s)L)^{-1}\lambda ds, \tag{36}$$

where $W_\lambda(t)$ is a standard Brownian motion under the measure $\mathbb{Q}$. Hence, we can define the temperature model under the measure $\mathbb{Q}$ as

$$dT(t) = dS(t) + \left[\lambda + \beta(t)\big(T(t) - S(t)\big)\right]dt + \sigma T(t)LdW(t). \tag{37}$$

3.4.5. Pricing CAT and GDD Futures on Temperature Basket

Definition 1. *For a a specificied contract period, $t \le t_1 < t_2$ and at a spatial location y_i, the CAT futures price is defined as in Equation (39):*

$$CAT_M[t_1, t_2] := \int_{t_1}^{t_2} M(t)$$
$$:= \sum_{i=1}^{N} \omega_i \left(\int_{t_1}^{t_2} T_i(x)dx\right). \tag{38}$$

From Equations (17), (38), and the linearity of expectation,

$$F_{CAT}(t, t_1, t_2; M) = \sum_{i=1}^{N} \omega_i \mathbb{E}_{\mathbb{Q}}\left[\int_{t_1}^{t_2} T_i(x)dx \mid \mathcal{F}_t\right]. \tag{39}$$

Definition 2. *For a a specificied contract period, $t \le t_1 < t_2$ and at location y_i, the CAT futures price is defined as*

$$GDD(\tau_1, \tau_2) := \int_{t_1}^{t_2} max\Big\{M(t) - C, 0\Big\}dx$$
$$= \int_{t_1}^{t_2} max\Big\{\sum_{i=1}^{N} \omega_i T_i(x) - C, 0\Big\}dx. \tag{40}$$

From Equations (2), (18), and using the linearity of expectation,

$$F_{GDD}(t, t_1, t_2; D) = \mathbb{E}_{\mathbb{Q}}\left(\int_{t_1}^{t_2} max\Big\{\sum_{i=1}^{N} \omega_i T_i(x) - C, 0\Big\}dx \Big| \mathcal{F}_t\right)$$
$$= \int_{t_1}^{t_2} \mathbb{E}_{\mathbb{Q}}\left(max\Big\{\sum_{i=1}^{N} \omega_i T_i(x) - C, 0\Big\} \Big| \mathcal{F}_t\right)dx. \tag{41}$$

Lemma 5. *If the dynamics of the daily average temperature follows Equation (37), then the explicit solution for the ith location y_i is given as*

$$T_i(t) = S_i(t) + \big(T_i(0) - S_i(0)\big)e^{\int_0^t \beta_i(s)ds} + \int_0^t \lambda_i e^{\int_u^t \beta_i(s)ds}du + \int_0^t \sigma_i T_i(u)e^{\int_u^t \beta_i(s)ds}\sum_{j=1}^{i} L_{ij}dW_j^\lambda(u).$$

Proof. The proof follows directly from the proof of Lemma 12 and observing the ith location y_i. □

Proposition 6. *At a spatial location i, the futures contract price on basket of CAT index following the mean-reverting regime in Equation (28) is calculated as*

$$F_{CAT}(t, t_1, t_2; M) = \sum_{i=1}^{N} \omega_i \left[\int_{t_1}^{t_2} S_i(x) + \int_{t_1}^{t_2} (T_i(t) - S_i(t)) e^{\int_t^x \beta_i(s)ds} + \right.$$

$$\int_{t}^{t_1} \int_{t_1}^{t_2} e^{\int_0^x \beta_i(s)ds} \lambda_i e^{\int_u^0 \beta_i(s)ds} dx du +$$

$$\left. \int_{u}^{t_2} \int_{t_1}^{t_2} e^{\int_0^x \beta_i(s)ds} \lambda_i e^{\int_u^0 \beta_i(s)ds} dx du \right].$$

Proof. For $x \geq t$ in Lemma 5,

$$T_i(x) = S_i(x) + (T_i(t) - S_i(t)) e^{\int_t^x \beta_i(s)ds} + \int_t^x \lambda_i e^{\int_u^x \beta_i(s)ds} du + \int_t^x \sigma_i T_i(u) e^{\int_u^x \beta_i(s)ds} \sum_{j=1}^{i} L_{ij} dW_j^\lambda(u) \quad (42)$$

$$\mathbb{E}_{\mathbb{Q}}\left[\int_{t_1}^{t_2} T_i(x)dx \mid \mathcal{F}_t \right] = \mathbb{E}_{\mathbb{Q}}\left[\int_{t_1}^{t_2} \left(S_i(x) + (T_i(t) - S_i(t)) e^{\int_t^x \beta_i(s)ds} + \int_t^x \lambda_i e^{\int_u^x \beta_i(s)ds} du + \right. \right.$$

$$\left. \left. \int_t^x \sigma_i T_i(u) e^{\int_u^x \beta_i(s)ds} \sum_{j=1}^{i} L_{ij} dW_j^\lambda(u) \right) dx \middle| \mathcal{F}_t \right]$$

$$= \int_{t_1}^{t_2} S_i(x)dx + \int_{t_1}^{t_2} (T_i(t) - S_i(t)) e^{\int_t^x \beta_i(s)ds} dx + \int_{t_1}^{t_2} \int_t^x \lambda_i e^{\int_u^x \beta_i(s)ds} du dx$$

$$= \int_{t_1}^{t_2} S_i(x) + \int_{t_1}^{t_2} (T_i(t) - S_i(t)) e^{\int_t^x \beta_i(s)ds} + L_1,$$

where

$$L_1 = \int_{t_1}^{t_2} \int_t^x \lambda_i e^{\int_u^x \beta_i(s)ds} du dx$$

$$= \int_{t_1}^{t_2} \int_t^{t_2} \mathbb{1}_{[t,x]}(u) \lambda_i e^{\int_u^x \beta_i(s)ds} du dx$$

$$= \int_{t}^{t_1} \int_{t_1}^{t_2} e^{\int_0^x \beta_i(s)ds} \lambda_i e^{\int_u^0 \beta_i(s)ds} dx du + \int_{u}^{t_2} \int_{t_1}^{t_2} e^{\int_0^x \beta_i(s)ds} \lambda_i e^{\int_u^0 \beta_i(s)ds} dx du$$

$$F_{CAT}(t, t_1, t_2; M) = \sum_{i=1}^{N} \omega_i \left[\int_{t_1}^{t_2} S_i(x) + \int_{t_1}^{t_2} (T_i(t) - S_i(t)) e^{\int_t^x \beta_i(s)ds} + \right.$$

$$\int_{t}^{t_1} \int_{t_1}^{t_2} e^{\int_0^x \beta_i(s)ds} \lambda_i e^{\int_u^0 \beta_i(s)ds} dx du +$$

$$\left. \int_{u}^{t_2} \int_{t_1}^{t_2} e^{\int_0^x \beta_i(s)ds} \lambda_i e^{\int_u^0 \beta_i(s)ds} dx du \right].$$

□

Proposition 7. *The GDD futures price for a contract time $t \leq t_1 < t_2$ at a given spatial location y^i for a basket of CAT index following the normal regime is given by*

$$F_{GDD}^N(t, t_1, t_2; D) = \int_{t_1}^{t_2} (\xi(t, x) + 2\Delta(t, x))^{\frac{1}{2}} (\phi(\Lambda(t, x)) + \Lambda(t, x)\Phi(\Lambda(t, x)))dx, \quad (43)$$

where Φ is the cumulative standard normal distribution function, ϕ is the standard normal density function,

$$\Lambda(t,x) = \frac{\psi(t,x) - C}{\left(\xi(t,x) + 2\Delta(t,x)\right)^{\frac{1}{2}}},$$

$$\psi(t,x) = \sum_{i=1}^{N} \omega^i \left(S_i(x) + \left(T_i(t) - S_i(t)\right)e^{\int_t^x \beta_i(s)ds} + \int_t^x \lambda_i e^{\int_u^x \beta_i(s)ds} du \right),$$

$$\xi(t,x) = \sum_{i=1}^{N} \omega_i^2 \sum_{j=1}^{i} \int_t^x \sigma_i^2 T_i(u)^2 e^{2\int_u^x \beta_i(s)ds} L_{ij}^2 du, \text{ and}$$

$$\Delta(t,x) = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \omega^i \omega^j \left(\sum_{q=1}^{i} L^{iq} L^{jq} \right) \int_t^x \sigma_i(u)\sigma_j(u)T_i(u)T_j(u)e^{\int_u^x (\beta_i(s)+\beta_j(s))ds} du$$

Proof. Let

$$D(x) = \sum_{i=1}^{N} \omega^i \tilde{T}_t^i. \tag{44}$$

For convenience, we denote the deterministic and random components of Equation (42) as $A_i(t,x)$ and $B_i(t,x)$, respectively. That is,

$$A_i(t,x) = S_i(x) + \left(T_i(t) - S_i(t)\right)e^{\int_t^x \beta_i(s)ds} + \int_t^x \lambda_i e^{\int_u^x \beta_i(s)ds} du,$$

$$B_i(t,x) = \int_t^x \sigma_i T_i(u)e^{\int_u^x \beta_i(s)ds} \sum_{j=1}^{i} L_{ij}dW_j^\lambda(u) = \sum_{j=1}^{i} \int_t^x \sigma_i T_i(u)e^{\int_u^x \beta_i(s)ds} L_{ij}dW_j^\lambda(u).$$

Hence,

$$D(x) = \sum_{i=1}^{N} \omega_i \left(A_i(t,x) + B_i(t,x) \right). \tag{45}$$

The distribution of the basket $D(x)$ at time t wll be established. $A_i(t,x)$ is, however, deterministic. By Itô isometry, and at t,

$$\int_t^x \sigma_i T_i(u)e^{\int_u^x \beta_i(s)ds} L_{ij}dW_j^\lambda(u) \sim N\left(0, \int_t^x \sigma_i^2 T_i(u)^2 e^{2\int_u^x \beta_i(s)ds} L_{ij}^2 du\right).$$

However, the $W_j^\lambda(u)$ are independent for each j. The variances can, therefore, be summed to obtain the variance of $B_i(t,x)$,

$$B_i(t,x) \sim N\left(0, \sum_{j=1}^{i} \int_t^x \sigma_i^2 T_i(u)^2 e^{2\int_u^x \beta_i(s)ds} L_{ij}^2 du\right) = N\left(0, \Psi^2(t,x)\right).$$

As $\sum_{i=1}^{N} \omega_i B^i(t,x)$ is a sum of normally distributed random variables, it is normally distributed with respective mean and variance:

$$\mathbb{E}\left(\sum_{i=1}^{N} \omega_i B^i(t,x)\right) = \sum_{i=1}^{N} \omega_i \mathbb{E}\left(B_i(t,x)\right) = 0,$$

and

$$Var\left(\sum_{i=1}^{N}\omega_i B_i(t,x)\right) = \sum_{i=1}^{N}Var\left(\omega_i B_i(t,x)\right) + 2\sum_{i<j}Cov\left(\omega_i H^i, \omega_j H^j\right)$$

$$= \sum_{i=1}^{N}\omega_i^2 Var(H_i(t,x)) + 2\sum_{i<j}\omega_i\omega_j Cov\left(H_i, H_j\right)$$

$$= \sum_{i=1}^{N}\omega_i^2 \Psi^2(t,x) + 2\sum_{i<j}\omega_i\omega_j Cov\left(B_i, B_j\right).$$

Consider $Cov\left(B_i, B_j\right)$ for $j > 1$. With respect to the standard Brownian motions $W_1^\lambda(u)$, both B_1 and B_j are in the same integral form. For $j > 1$ and by the independence of $W_1^\lambda(u)$ and $W_j^\lambda(u)$, the covariance only exists between these two integrals. Therefore,

$$Cov\left(B_i, B_j\right) = \int_t^x \sigma_i(u)\sigma_j(u)T_i(u)T_j(u)e^{\int_u^x(\beta_i(s)+\beta_j(s))ds}\left(\sum_{q=1}^{i}L^{iq}L^{jq}\right)du, \quad \forall j > 1$$

$$= \left(\sum_{q=1}^{i}L^{iq}L^{jq}\right)\int_t^x \sigma_i(u)\sigma_j(u)T_i(u)T_j(u)e^{\int_u^x(\beta_i(s)+\beta_j(s))ds}du, \quad \forall j > 1.$$

Define $\mathrm{Y}ij(t,x) := \int_t^x \sigma_i(u)\sigma_j(u)T_i(u)T_j(u)e^{\int_u^x(\beta_i(s)+\beta_j(s))ds}du,$

$$Var\left(\sum_{i=1}^{N}\omega_i B_i(t,x)\right) = \sum_{i=1}^{N}\omega_i^2\Psi^2(t,x) + 2\sum_{i<j}\omega_i\omega_j\left(\sum_{q=1}^{i}L_{iq}L_{jq}\right)\mathrm{Y}ij(t,x)$$

$$= \sum_{i=1}^{N}\omega_i^2\Psi^2(t,x) + 2\sum_{i=1}^{N}\sum_{j=i+1}^{N}\omega_i\omega_j\left(\sum_{q=1}^{i}L_{iq}L_{jq}\right)\mathrm{Y}_{ij}(t,x).$$

From (45),

$$D(x) \sim N\left(\sum_{i=1}^{N}\omega_i A_i(t,x), \sum_{i=1}^{N}\omega_i^2\Psi^2(t,x) + 2\sum_{i=1}^{N}\sum_{j=i+1}^{N}\omega_i\omega_j\left(\sum_{q=1}^{i}L_{iq}L_{jq}\right)\mathrm{Y}_{ij}(t,x)\right).$$

Let

$$\psi(t,x) = \sum_{i=1}^{N}\omega_i A_i(t,x); \quad \xi(t,x) = \sum_{i=1}^{N}\omega_i^2\Psi^2(t,x);$$

$$\Delta(t,x) = \sum_{i=1}^{N}\sum_{j=i+1}^{N}\omega_i\omega_j\left(\sum_{q=1}^{i}L_{iq}L_{jq}\right)\mathrm{Y}_{ij}(t,x).$$

$D(x)$ can be written, in the form of a standard normal random variable $Z \sim N(0,1)$, as

$$D(x) = \psi(t,x) + \left(\xi(t,x) + 2\Delta(t,x)\right)^{\frac{1}{2}}Z. \tag{46}$$

Recall Equation (41) and consider

$$\sum_{i=1}^{N} \omega_i T_i(x) - C > 0,$$

$$\left(\xi(t,x) + 2\Delta(t,x) \right)^{\frac{1}{2}} Z > C - \psi(t,x),$$

$$Z > \frac{C - \psi(t,x)}{\left(\xi(t,x) + 2\Delta(t,x) \right)^{\frac{1}{2}}} := \Lambda^1(t,x).$$

From the above equation,

$$C = \psi(t,x) + \Lambda^1(t,x) \left(\xi(t,x) + 2\Delta(t,x) \right)^{\frac{1}{2}}. \tag{47}$$

It can be deduced from, Equations (41) and (47), that

$$\mathbb{E}_{\mathbb{Q}} \left(max \left\{ \sum_{i=1}^{N} \omega_i T_i(x) - C, 0 \right\} dx \Big| \mathcal{F}_t \right) = \int_{\Lambda^1(t,x)}^{+\infty} \left(D(x) - C \right) \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}} dz. \tag{48}$$

Then,

$$\mathbb{E}_{\mathbb{Q}} \left(max \left\{ \sum_{i=1}^{N} \omega^i \tilde{T}_t^i - K, 0 \right\} dx \Big| \mathcal{F}_t \right) = \int_{\Lambda^1(t,x)}^{+\infty} \left(\psi(t,x) + (\xi(t,x) + 2\Delta(t,x))^{\frac{1}{2}} z - \right.$$

$$\psi(t,x) - \Lambda^1(t,x) \left(\xi(t,x) + 2\Delta(t,x) \right)^{\frac{1}{2}} \right) \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}} dz$$

$$= \int_{\Lambda^1(t,x)}^{+\infty} \left((\xi(t,x) + 2\Delta(t,x))^{\frac{1}{2}} z - \right.$$

$$\Lambda^1(t,x) \left(\xi(t,x) + 2\Delta(t,x) \right)^{\frac{1}{2}} \right) \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}} dz$$

$$= \left(\xi(t,x) + 2\Delta(t,x) \right)^{\frac{1}{2}} \left(\int_{\Lambda^1(t,x)}^{+\infty} \frac{z e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}} dz + \Lambda^1(t,x) \Phi \left(-\Lambda^1(t,x) \right) \right)$$

$$= \left(\xi(t,x) + 2\Delta(t,x) \right)^{\frac{1}{2}} \left(\frac{e^{-\frac{1}{2}\Lambda(t,x)^2}}{\sqrt{2\pi}} + \Lambda(t,x) \Phi \left(\Lambda(t,x) \right) \right)$$

$$= \left(\xi(t,x) + 2\Delta(t,x) \right)^{\frac{1}{2}} \left(\phi \left(\Lambda(t,x) \right) + \Lambda(t,x) \Phi \left(\Lambda(t,x) \right) \right).$$

Therefore,

$$F_{GDD}^N(t,t_1,t_2;D) = \int_{t_1}^{t_2} \left(\xi(t,x) + 2\Delta(t,x) \right)^{\frac{1}{2}} \left(\phi \left(\Lambda(t,x) \right) + \Lambda(t,x) \Phi \left(\Lambda(t,x) \right) \right) dx,$$

where

$$\Lambda(t,x) = -\Lambda^1(t,x) = \frac{\psi(t,x) - C}{\left(\xi(t,x) + 2\Delta(t,x) \right)^{\frac{1}{2}}}.$$

$\square$

4. Discussion and Conclusion

The agricultural sector employs a large workforce in Ghana and serves as the principal source of income for most people, especially small-holder farmers in the Northern savanna. This sector is vulnerable to climate shocks and, hence, there is a need for a reliable and efficient insurance product (weather derivative) for small-holder farmers and stakeholders. However, most farmers are unwilling to buy this product as a result of high basis risks in the product design and pricing. To reduce basis risk in weather derivative design and pricing, in this study, the historical relationship between maize yield and some selected weather variables were determined by using machine learning ensemble technique and feature importance. Maize yield was chosen as the proxy for crop yields, due to its economic importance to the farmers in the Northern savanna. The feature importance gave a score of the importance of each weather variable in building the ensemble-learning model. The results indicated that average temperature was the most important weather variable that affected the yield of maize. Consequently, average daily temperature was used as the underlying weather variable for weather derivative design. In this way, product design basis risk was mitigated.

A time-varying daily average temperature model was proposed. The model was a mean-reverting process with time-varying speed of mean-reversion, seasonal mean, and a local volatility that captured the local variations of the daily average temperature. The model was extended to a multi-dimensional model for different but correlated locations. For analytical tractability for the pricing models, the residuals of the daily average temperature were assumed to be normally distributed. Our models captured most of the stylized facts of temperature, such as mean-reversion, seasonality, volatility, and locality features of the selected locations. Using the proposed average temmodels, closed-form pricing formulas for futures, options on futures, basket futures for cumulative average temperature (CAT), and growing degree-days (GDD) were presented. Since there is not yet a real WD market for the selected locations, we assume a constant market price of risk in the pricing models.

With these efficient and reliable pricing models, basis risks will be mitigated. As a result, there will be an increase in the willingness to pay for the contracts on the farmers' side and trading activities in the WD market will also increase. Using the proposed spatial-temporal pricing model, it will be cost efficient to buy contracts for different but correlated farming locations rather than a single farming location. Farmers and other agricultural stakeholders can control covariate risks and hedge their crops against the perils of weather uncertainties.

In summary, basis risks (product-design and geographical) can mitigated if weather derivatives are properly designed.

References

1. UNECA. Challenges to agricultural development in Africa. In *Economic Report on Africa 2009 Developing African Agriculture Through Regional Value Chains*; UNECA: Addis Ababa, Ethiopia, 2009. Available online: https://www.uneca.org/sites/default/files/PublicationFiles/era2009_eng_full.pdf (accessed on 18 July 2019).
2. Ibn Musah, A.A.; Du, J.; Bilaliib Udimal, T.; Abubakari Sadick, M. The nexus of weather extremes to agriculture production indexes and the future risk in Ghana. *Climate* **2018**, *6*, 86. [CrossRef]
3. Turvey, C.G. Weather derivatives for specific event risks in agriculture. *Rev. Agric. Econ.* **2001**, *23*, 333–351. [CrossRef]

4. Zong, L.; Ender, M. Spatially-aggregated temperature derivatives: Agricultural risk management in China. *Int. J. Financial Stud.* **2016**, *4*, 17. [CrossRef]

5. Woodard, J.D.; Garcia, P. Basis risk and weather hedging effectiveness. *Agric. Finance Rev.* **2008**, *68*, 99–117. [CrossRef]

6. Rohrer, M. The relevance of basis risk in the weather derivatives market. *WIT Trans. Model. Simul.* **2004**, *38*. [CrossRef]

7. Ritter, M.; Musshoff, O.; Odening, M. Minimizing geographical basis risk of weather derivatives using a multi-site rainfall model. *Comput. Econ.* **2014**, *44*, 67–86. [CrossRef]

8. Spicka, J.; Hnilica, J. A methodical approach to design and valuation of weather derivatives in agriculture. *Adv. Meteorol.* **2013**, *2013*, 146036. [CrossRef]

9. Musshoff, O.; Odening, M.; Xu, W. Management of climate risks in agriculture—Will weather derivatives permeate? *Appl. Econ.* **2011**, *43*, 1067–1077. [CrossRef]

10. Zariphopoulou, T. A solution approach to valuation with unhedgeable risks. *Finance Stoch.* **2001**, *5*, 61–82. [CrossRef]

11. Henderson, V. Valuation of claims on nontraded assets using utility maximization. *Math. Finance* **2002**, *12*, 351–373. [CrossRef]

12. Owen, M.P. Utility based optimal hedging in incomplete markets. *Ann. Appl. Probab.* **2002**, *12*, 691–709. [CrossRef]

13. Davis, M. Pricing weather derivatives by marginal value. *Quant. Finance* **2001**, *1*, 305–308. [CrossRef]

14. Brockett, P.L.; Goldens, L.L.; Wen, M.-M.; Yang, C.C. Pricing weather derivatives using the indifference pricing approach. *N. Am. Actuar. J.* **2009**, *13*, 303–315. [CrossRef]

15. Barrieu, P.; Karoui, N.E. Optimal design of derivatives in illiquid markets*. *Quant. Finance* **2002**, *2*, 181–188. [CrossRef]

16. Cao, M.; Wei, J. *Equilibrium Valuation of Weather Derivatives*; Queens University: Kingston, ON, Canada, 2000.

17. Lucas, R.E., Jr. Asset prices in an exchange economy. *Econom. J. Econom. Soc.* **1978**, *46*, 1429–1445. [CrossRef]

18. Alaton, P.; Djehiche, B.; Stillberger, D. On modelling and pricing weather derivatives. *Appl. Math. Finance* **2002**, *9*, 1–20. [CrossRef]

19. Mraoua, M. Temperature stochastic modeling and weather derivatives pricing: Empirical study with Moroccan data. *Afrika Statistika* **2007**, *2*, 22–43. [CrossRef]

20. Benth, F.E.; Benth, J.Š. Stochastic modelling of temperature variations with a view towards weather derivatives. *Appl. Math. Finance* **2005**, *12*, 53–85. [CrossRef]

21. Benth, F.E.; Benth, J.Š. The volatility of temperature and pricing of weather derivatives. *Quant. Finance* **2007**, *7*, 553–561. [CrossRef]

22. Hu, Q.; Buyanovsky, G. Climate effects on corn yield in Missouri. *J. Appl. Meteorol.* **2003**, *42*, 1626–1635. [CrossRef]

23. Wood, T.N. Agricultural development in the Northern Savannah of Ghana. Ph.D. Thesis, University of Nebraska, Lincoln, NE, USA, 2013.

24. Alexandridis, A.; Zapranis, A.D. *Weather Derivatives: Modeling and Pricing Weather-Related Risk*; Springer Science & Business Media: New York, NY, USA, 2012.

25. Jones, C.S. The dynamics of stochastic volatility: Evidence from underlying and options markets. *J. Econom.* **2003**, *116*, 181–224. [CrossRef]